An Efficient Algorithm for Interested Object Detection in the Video Sequences

Kanagamalliga. S[1*], Santhiya Bhavani. K[2], Shobana. V[3], Sundara Priya. T[4]

[1]Assistant Professor, Department of Electronics and Communication Engineering, Madurai-625009.

[2,3,4]UG Student, Department of Electronics and Communication Engineering, Madurai-625009.

Abstract— **A scale invariant feature transform (SIFT) is used for object tracking in real scenarios. SIFT features are used to correspond the region of interests across the frames. Experimental work demonstrates that the proposed SIFT strategy improves the tracking performance than the classical SIFT tracking algorithms in complicated real scenarios. This paper briefs about tracking selected objects by both single and multiple tracking methods using region based tracking. As far as humans are concerned object classification in video is a simple task but it is a complex and challenging task for machines due to different factors such as object size, occlusion, scaling, etc**. **The features calculated from the frames to be registered must be distinctive and then it can be matched. Feature based clustering can be used in object recognition and video tracking**. **By conducting an experimental study on a special dataset, a commitment can be made that the proposed method provides the satisfactory results by achieving an overall accuracy rate.**

Keywords: *Video data, Frame conversion, Object detection, Object tracking, Scale Invariant Feature Transform.*

I. INTRODUCTION

Object tracking is a very important topic in multimedia technologies, particularly in applications such as teleconferencing, surveillance and human–computer interface [1].The goal of object tracking is to determine the position and to track the object in videos continuously and reliably against static and dynamic scenes [2]-[4]. To achieve this target, a number of algorithms have been established [5]. In the situations, additional features may be used as a complement that can improve the capability of the trackers [6]. Unfortunately, edges, corners and silhouette are application based features, and cannot effectively work in case of variable scaled, rotated or translated images [7].

In recent years, some of the decent techniques have been used for the purpose of optimal solutions to these problems [8]. For example, scale invariant feature transform was used to generate feature points in a full image [9]. These feature points appear invariant for any scaling, rotation or translation of the images [11]. Therefore, the SIFT features can be used with an established tracking system for improving the performance of the latter [16].

In this paper the proposed tracking algorithm is an effective SIFT feature tracking. The proposed between two neigh-boring frames in terms of color and SIFT correspondence [10]. Technically, a track will be made if SIFT feature tracking lead to approximate probability

distributions within the corresponding region in the next image frame [13]-[15]. This algorithm is employed in order to pursue a maximum estimation using the measurements from SIFT correspondence [14]. The main contributions of this paper consist of the tracking performance of the proposed strategy can be experimentally revealed using SIFT for better accuracy level and the quality of the images.

II. PROPOSED METHOD

Fig. 1. Shows the flow of the proposed method. The input given is in video format. For further process the video is converted into frames. From the converted frames, the features are extracted using SIFT. This feature extraction can be performed with the help of SIFT features. After the completion of the feature extraction, region based selection is performed.

Then the interested object will be detected. After that, the selected or interested region will be tracked. Performance measure is the last step in this process which reveals overall performance based on the accuracy and the efficiency of the system.
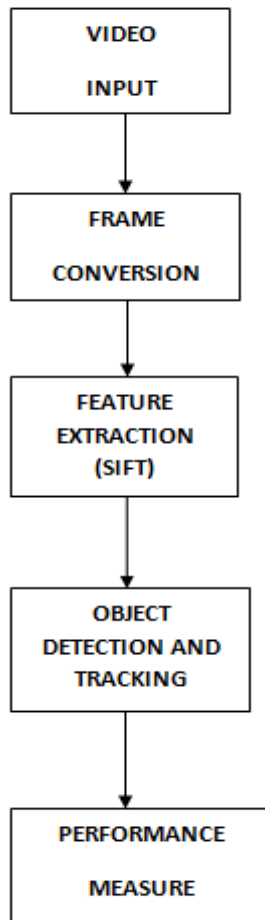
```
┌─────────────┐
│   VIDEO     │
│   INPUT     │
└─────────────┘
       │
       ▼
┌─────────────┐
│   FRAME     │
│ CONVERSION  │
└─────────────┘
       │
       ▼
┌─────────────┐
│  FEATURE    │
│ EXTRACTION  │
│   (SIFT)    │
└─────────────┘
       │
       ▼
┌─────────────┐
│   OBJECT    │
│ DETECTION AND│
│  TRACKING   │
└─────────────┘
       │
       ▼
┌─────────────┐
│ PERFORMANCE │
│  MEASURE    │
└─────────────┘
```

Fig. 1. Flow diagram of the proposed method

A. VIDEO INPUT

Fig. 2. is the video signal given as input. A video input is a port or a jack is that receives a video signal from one device to other device. A Video file is normally consists of a container format containing video coding format alongside audio data in audio format. There are many types of video formats. For example, AVI (Audio Video Interleave), FLV (Flash Video), WMV (Windows Media Video), MP4 (Moving Pictures Expert Group 4), and their file extensions are .avi, .fly, .wmv, .mp4 respectively. The most popular video format is .avi.

B. FRAME   CONVERSION

Fig. 3. Shows the conversion of video input into frames. Frames can be obtained from a video and converted into images. The required numbers of frames are attained from a video. The conversion from a video into frames is the second step.

The Frame Conversion block passes the input, and the output is taken with the set of output sampling mode to the value of the Sampling mode of output signal parameter, which can be either Frame-based or Sample-based. The output sampling mode can also be inherited from the signal at the reference input port

C. FEATURE EXTRACTION

Feature extraction starts from a set of measured data and builds derived features intended to be informative and non-reduntant, facilitating the subsequent learning and generalization steps. Feature extraction is related to dimensionality reduction which is useful when image size are large and a reduced feature representation is required to quickly complete tasks such as image matching and retrieval. The low-level detections available are: Edge detection, Corner detection, Blob detection, Ridge detection, SIFT.

Feature extraction is done by SIFT, to detect and describe local features in images. Some of the applications of SIFT include object recognition, robotic mapping and navigation. Each of the SIFT key points specifies the 2D location, scale, orientation, and each matched key point in the database has a record of its parameters relative to the training image in which it was found.

The evaluation carries strongly that SIFT-based descriptors, which are region-based, are the most robust and characteristic, and are therefore best suited for feature matching. The performance of frames matching the SIFT descriptors can be improved in the sense of achieving higher efficiency.

D. OBJECT DETECTION AND TRACKING

Object tracking is for the purpose of tracking the movement of the human. Object detection algorithms typically used to extract the features and learning algorithms to recognize instances of an object category. Object is also used for identifying a specific object in a digital image or video. It relies mainly on matching or pattern recognition algorithms using appearance-based or feature-based techniques.

The object detection requires more than simple region appearance features. The object detector's ability is to localize appearance and general shape characteristics of a class. Thus, in addition to raw appearance features, it appends to object features are derived from object detection.

The experiments test the approach which learns feature vectors constructed by computing histograms of gradient orientations in fixed-size overlapping cells within the candidate window. A better approach is to mask out all pixels not belonging to the object. The soft mask will make intensity of pixels outside the object based on their distance to the object boundary. It shows the dual advantage of preventing hard edge artifacts and being less sensitive to segmentation errors. The masked window is used at both training and test time. Object appearance and shape are captured by operating on both the original image and the edge-filtered image. Object detection can be done by both approaches.

K-means clustering aims to partition n observations into k clusters where observation belongs to the cluster with the nearest mean, as a prototype of the cluster. The K-means clustering algorithm clusters data by repeatedly computing a mean intensity for each class and segmenting

Special Issue - 2017

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCIECC - 2017 Conference Proceedings

the image by classifying each and every pixel in the class with the closest mean. This clustering algorithm is an unsupervised algorithm, it is used to segment the interest area from the background. But before applying K -means algorithm, first partial stretching enhancement is applied to the image for improving the quality of the image. Subtractive clustering method is data clustering method which generates the centroid based on the potential value of the data points.

III. EXPERIMENTAL RESULTS

To validate the proposed method, a variety of experiments are conducted on a special dataset and the performance of proposed method. The test input which has the presence of human, objects on it. The videos are converted into 15 frames per second and each frame is analyzed.

Fig.1 shows that, from the static video which is used as an input.Fig.2 reveals that the input video is converted into frames .The total frames obtained from the video is 110.From the frames attained, frame number 01 and 110 is taken as samples.Fig.3 shows that the obtained frame conversion with frame number 110 is taken as the sample.

The feature is extracted for the frame number01 as shown in Fig.4.Fig.5 shows that the feature extraction which maps the image pixels into feature space. Both Fig.4 and Fig.5 reveals the process of Feature Extraction.Fig.6 involves the process of interested object detection.
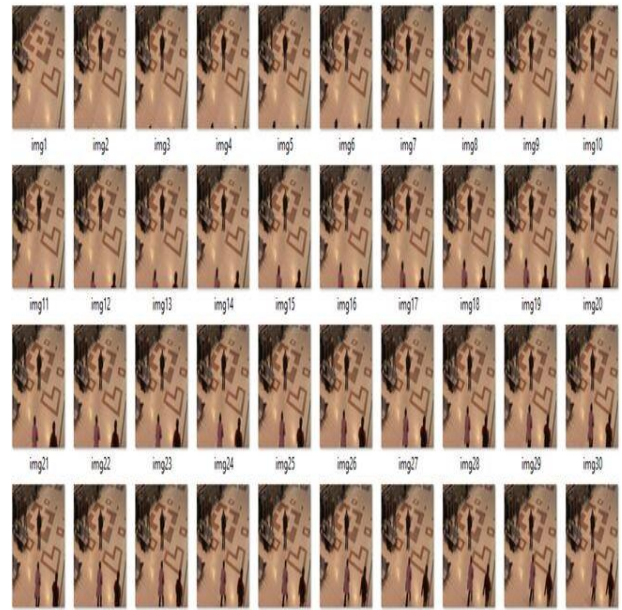

Fig. 2. Frame conversion for input video signal

Fig. 3. Input video signal (Frame number 110)

Fig. 1. Video signal given as an input (Frame number 01)

Fig. 4. Feature extraction (Frame number 01)

Special Issue - 2017

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
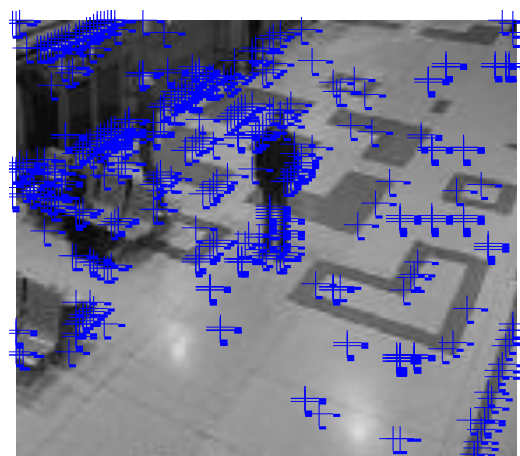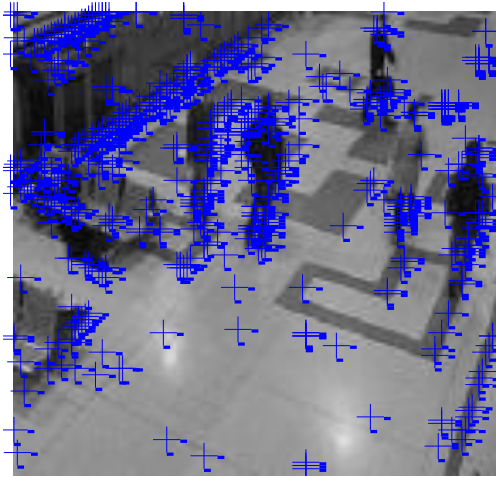**NCIECC - 2017 Conference Proceedings**

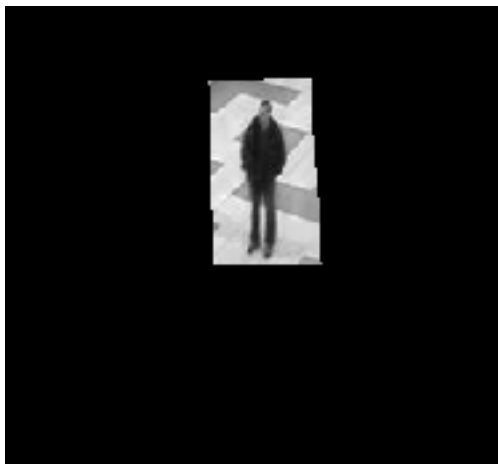Fig. 5. Feature extraction (Frame number 110)

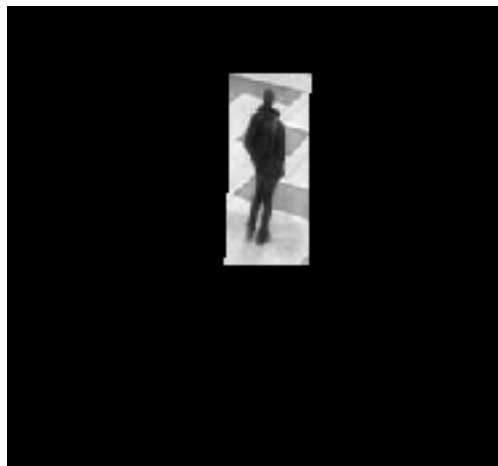Fig. 6. Interested object detection (Frame number 01)

Fig. 7. Interested object detection (Frame number 110)

Object is detected by the difference between the current frame and a reference frame, called as the background image. The variations between current video frames to that of the reference frame in terms of pixels signify existence of moving objects. Object detection is a simple algorithm and is highly sensitive to identify object. Fig.7 implies that the frame number 110 which is the detection of an interested object. As there is a movement of more than two humans as shown in Fig.1, this paper proposes to detect interested object. Hence Fig.6 and Fig.7 shows an interested human is being tracked for both frame number 01 and frame number 110.

IV. CONCLUSION

This paper deals with a robust approach towards object detection and tracking. Main focus of the presented work is to effectively locate the object by using SIFT descriptor. Tracking is applied for both single and multiple interested objects. A solution to enhance the performance of SIFT using object tracking has been presented. This work integrated the outcomes of SIFT feature correspondence and its tracking and it could produce better solution for object tracking in different scenarios. The experimental results obtained on the dataset; shows that the utilized features are discriminative and robust when the contribution is considered on the basis of accuracy rate. The result of this algorithm is applicable in all types of cases. The overall performance of the proposed technique achieves better accuracy than other techniques.

REFERENCES

[1] Ji Zhao, Liantao Wang, Ricardo Cabral Features and Region Selection for Visual learning,2016.

[2] Murat Olgun, Ahmet Okan Onarcan, Kemal Özkan , Wheat Grain Classification by using D-SIFT Features with SVM Classifier,2016.

[3] A.Azzem, M.sharif, J.H.Shah, Hexagonal scale invariant feature transform for facial feature extraction,2015.

[4] .Tarek Elguebaly a, Nizar Bouguila b, Simultaneous high-dimensional clustering and feature selection using asymmetric Gaussian mixture models 2015.

[5] Jalal A,Uddin M Z, Kim T–S, Depth video-based human activity recognition system using translation and scaling invariant features for life logging at smart home. Consumer Electronics; IEEE Transactions on, vol.58, no.3, pp.863, 871; August 2012.

[6] Narhe, MeghaC, and M.S.Nagmode, Vehicle Classification using SIFT. International Journal of Engineering Research and Technology.Vol. 3. No. 6 2014. ESRSA Publications, 2014.

[7] Sun, L., Ji, S., & Ye, J. 2011. Canonical correlation analysis for multilabel classification: A least-squares formulation, extensions, and analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33, 194-200.

[8] R.P. Browne, P.D. McNicholas, M.D. Sparling, Model-based learning using a mixture of mixtures of Gaussian and uniform distributions, IEEE Trans. Pattern Anal. Mach. Intell. 34 2012 814–817.

[9] Sharif, M., Ayub, K., Sattar, D., Raza, M., & Mohsin, S. 2012. Enhanced and Fast Face Recognition by Hashing Algorithm.

[10] X. Liu, L. Wang, J. Yin, and L. Liu, "Incorporation of radius-info can be simple with Simple MKL," Neurocomputing, vol. 89, pp. 30–38,Jul. 2012.

[11] O. Yakhnenko, J. Verbeek, and C. Schmid, "Region-based image classification with a latent SVM model," INRIA, Rocquencourt, France, Tech. Rep. RR-7665, 2011.

[12] M. Kloft, U. Brefeld, S. Sonnenburg, and A. Zien, "p-norm multiple kernel learning," J. Mach. Learn. Res., vol. 12, pp. 953–997, Mar. 2011.

[13] A.Azzem, M.sharif, J.H.Shah, Hexagonal scale invariant feature transform for feature extraction,2015.

Special Issue - 2017

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCIECC - 2017 Conference Proceedings

[14]   Y. Cheng, SIFT, mode seeking, and clustering, IEEE Trans. Pattern Anal.Mach. Intell 2013.

[15]   D. Comaniciu, P. Meer, SIFT: a robust approach toward feature space analysis, IEEE Trans. Pattern Anal. Mach. Intell. 2010.

[16]   M. Fashing, C. Tomasi,SIFT is a bound optimization, IEEE Trans. Pattern Anal. Mach. Intell.2011.